DOCUMENT RESUME

ED 221 594 TM 820 632

AUTHOR Chafin, Ann E.; Lindheim, Elaine
TITLE Standards for Standard Setting.
PUB DATE Mar 82
NOTE 21p.; Paper presented at the Annual Meeting of the American Educational Research Association (66th, New York, NY, March 19-23, 1982).

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS *Cutting Scores; Data Analysis; *Graduation Requirements; High Schools; *Minimum Competency Testing; Performance; *Quantitative Tests; Secondary School Mathematics; *State Standards; Student Certification; Test Items
IDENTIFIERS Domain Specifications; *Maryland Functional Mathematics Test; Maryland State Department of Education; *Standard Setting

ABSTRACT

The procedures projected to be used in the spring and summer of 1982 to establish passing standards for the Maryland Functional Mathematics Test are described. The development of the test as a prerequisite for high school graduation to be implemented in 1987 is described. The maximum involvement of diverse constituencies and multiple informed judgments by the Maryland State Board of Education are requirements in the procedures for test item-based judgments, test specification-based judgments, and use of student performance data by standard-setters. The information gathering plan provides for sessions on item-based judgments with selected participants for school districts using an information pooling system in a prepared script to insure administrative consistency. Teachers, parents and students will be sampled by mail with a test description and domain item samples to develop specification-based judgments. During the fall of 1982 a cut-score validity study of ninth grade tests will be conducted. Multiple recommended cut scores will be used in making a decision and a Board of Education bylaw on performance standards. (CM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED221594

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY
A. E. Chafin
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

STANDARDS FOR STANDARD SETTING*

by

Ann E. Chafin

Maryland State Department of Education

and

Elaine Lindheim

IOX Assessment Associates

According to a projected implementation calendar, beginning with the graduating class of 1987, Maryland high school students will need to pass the Maryland Functional Mathematics Test (MFMT) in order to be awarded a high school diploma. The creation of a functional mathematics test as a graduation prerequisite has now reached the fourth major phase of its development, namely, the setting of a performance standard. The following paper describes how the Maryland State Department of Education (MSDE) will conduct this standard-setting enterprise. However, because the standard that is set on the MFMT can be no more defensible than the assessment instrument upon which it is based, it is appropriate to precede the description of standard-setting with a brief history of the development of the MFMT.

Development of the mathematics test began in 1978, with the identification of content areas that Maryland educators, parents,

---

* A symposium presentation at the annual meeting of the American Educational Research Association, New York, March 19-23, 1982.

TM 820 632

and citizens felt were appropriate for coverage by a functional mathematics program. These content areas were adopted by the Maryland State Board of Education and subsequently disseminated throughout the state in the MSDE publication, Declared Competencies Index (1980).

The second step in the test development process took place in March of 1981 when a group of Maryland mathematics teachers and curriculum specialists used the 30 objectives in the Declared Competencies Index to develop test specifications for an assessment system. Seven different assessment domains were identified and the rules for generating test items to measure each domain carefully delineated. The resulting document (Test Specifications for the Maryland Functional Mathematics Test, Maryland State Department of Education, Baltimore, November 1981) was used to guide the production of 546 mathematics test items, 78 items per assessment domain. Half of these items were developed for use in the upper level version of the MFMT. This form of the measure is given for the first time in the ninth grade and used for graduation certification. The other half of the test items were developed for use in the lower level version of the MFMT, a seventh-grade examination that serves a diagnostic and instructional function.

In the fall of 1981, all of the mathematics test items were field tested with representative samples of Maryland students. A total

of 24,000 students took part in this field test, including 9,000 ninth-graders and 3,000 eleventh-graders who participated in the tryout of the upper level MFMT items. Data analyses from this field test provided various indices of item quality which were subsequently used to guide the construction of the first operational forms of the MFMT. The field test results also yielded student performance data. With an operational test form in place, MSDE now must proceed to gather information that will help guide the State Board of Education as it makes its important decision regarding an appropriate passing standard.

Criteria for Standard-Setting

Establishing the passing standard for a certification examination can be the most vulnerable operation in the entire development of that instrument. Since the introduction and widespread use of minimum competency tests in the last few years, the issue of cut-score determination has received a great deal of attention. Two key issues seem to emerge in all such considerations. First, standard-setting involves political as well as psychometric concerns. To be credible, a standard-setting operation should involve as many different concerned constituencies as practically possible. The more individuals who participate in the process of recommending a cut-score, the more credibility that process will possess.

Second, standard-setting is at heart a judgmental operation. Although there are many different procedures for standard-setting

currently in use, all of them rely at some point on human judgment. Judgmentally-based decisions, if not made carefully, will always be subject to the criticism that they are capricious. One strategy for avoiding such charges is to provide the individual or group who must make the final decision regarding a cut-score with as much information as possible. Multiple sources of information hopefully will allow for the most informed judgment possible.

The standard-setting procedures to be used for the Maryland mathematics tests have been designed to meet the two key criteria of maximum participation and multiple types of information. The question of "How well must a student do in order to pass the Maryland Functional Mathematics Tests?" will ultimately be a decision of the State Board of Education. These decision-makers will be provided with information of three different types, namely, test item-based judgments, test specifications-based judgments, and student performance data. The following pages elaborate on how each type of information will be collected.

Item-Based Judgments

The MSDE has chosen an item-based standard-setting method similar to the Modified Delphi procedure used by Dr. Richard Jaeger to investigate the passing standards for the North Carolina High School Competency Tests in Reading and Mathematics (Jaeger 1981). Some major departures from the procedures outlined by Jaeger will be made to facilitate the completion of such an activity in Maryland.

Rather than being sampled, participants will be invited by district superintendents, with each of the twenty-four districts asked to select a total of six individuals; four citizens, one mathematics teacher, and one administrator. This deviation from Jaeger's procedure was chosen as a measure to insure a higher participation rate.

The meetings will be held simultaneously at three sites in the state on a Saturday in the spring of this year. The information collected at each site will be telephoned to a central location, where it will be pooled and then communicated back to each meeting site. Because of Maryland's particular geography it is possible that as few as three districts could be represented at one of the regional locations. Using the electronic pooling of data offers more anonymity for the judges at each site.

The sessions will be conducted using a prepared script to insure administrative consistency. After a preliminary orientation, judges will be asked to answer the eighty mathematics test items as examinees, marking their responses on an answer sheet. (At the end of the day, raters will be given a scoring key and allowed to score their responses, if they desire. All test keys and answer sheets, as well as all test items, will be collected at the end of the session.) After they have taken the test, judges will begin their standard-setting procedure.

Judges will be asked to consider each test item and to answer "Yes" or "No" to the question, "Should every regular high school graduate in Maryland be able to answer this item correctly?" Judges will be told, "If you answer 'Yes' for an item, you are saying that it tests something that should be required of every regular student who is awarded a high school diploma. If you answer 'No' for an item, you are saying that it tests something that should not be required of every regular student who is awarded a high school diploma."

After the first round of balloting has been completed and tabulated, each judge's response sheet will be returned with the mean score recommended by their group of participants recorded in the corner. Judges will then be asked to make any adjustment to their recommended cut-score they feel is necessary in the light of this peer feedback. This information will not be collected until after the final round of balloting.

As soon as the second ballot is complete, judges will open sealed packets containing student performance data for ninth and eleventh grade students. This data will consist of each item's difficulty for the two grade levels separately and the percentage of ninth and eleventh graders who would pass the test at each of a range of possible pass scores. Based on this performance data participants will be asked to make a final judgment about each of the eighty items.

The session will conclude with a question and answer period during which time participants can discuss any concerns they might have about the testing program. Emerging from these meetings will be nine recommended passing scores, one from each group of participants, after each round.

Specifications-Based Judgments

Mailing to Maryland Schools. Maryland's 312 secondary schools will be invited, by mail, to have teachers, parents, and students offer advice regarding the mathematics test performance standard. Each school will receive packets of materials to distribute to the three groups. Samples of these materials are appended. These materials will include letters to participants and an outline of the MFMT which provides a test description and a sample item for each domain tested. Directions for sampling participants and completing the rating task, a form for summarizing the school's preferences (in the form of the percent correct score recommended by each group and the number of respondents in each group) and a stamped, pre-addressed envelope in which to return the rating form will also be included in each school's mailing. Half of the schools will receive a summary of student performance data, to be used by the teachers at that school in making their judgments about a performance standard. The letter to each principal will indicate that a school's participation in this procedure is entirely voluntary and that all responses will be strictly confidential.

Blue Ribbon Reactors. Selected Maryland citizens (for example, individuals involved with personnel selection at Maryland-based business concerns, college mathematics professors and business executives) will be sent an invitation to provide a recommended standard for the MFMT. Two separate mailings will be developed, each to be sent to about 100 respondents. Both mailings will contain test description and sample items. One of the mailings will also contain student performance data. All of the reviewers will be asked to recommend the percentage of items that a regular Maryland high school graduate should answer correctly in order to pass.

Student Performance Data

The performance data presented will allow standard-setters to review student performance by grade level (with ninth and eleventh grade performance, but not seventh grade included) and also to see the proportions of students who would fail the test at varying passing standards. This information will be presented to all participants in the item-based procedure, half of the schools in the mailing to Maryland high schools, and half of the Blue Ribbon Reactors.

Creation of a Functional Mathematics Bylaw

Before the Maryland Functional Mathematics Tests can become a graduation prerequisite, the Maryland State Board of Education must pass a bylaw stating the standard for passing and the first graduating class (probably 1987) that must meet the requirement.

At the end of the procedures outlined above, there will be multiple recommended cut-scores, one each from the various constituency groups.

To proceed from several recommended cut-scores to a bylaw requires some synthesis of the information about the recommendations and student performance. Although this synthesis could occur during State Board discussions, it is more likely that the Division of Instruction will go to the State Superintendent with one recommended score, a rationale to back up the score and estimated performance data.

Cut-Score Validity Study for Ninth-Grade Tests

After a standard is chosen, the validity of this cut-score will be estimated during the fall of 1982. Two groups of examinees, identified as masters and non-masters, will be formed on the basis of teacher judgment. The "charge" to teacher judges will require particular attention. The precise form of that request for judgment can result in substantially different ways of classifying masters/non-masters. Once the master and non-master groups are formed and the tests are administered, a number of different indices will be used to assess the validity of the cut-score. The first of these indices is the probability of making a correct decision (i.e., the proportion of accurate decisions, namely, the proportion of agreement between decisions based on test scores and decisions based on teacher judgment).

Another is the ratio of errors in classification type I/type II (i.e., the proportion of true masters falsely declared to be non-masters on the basis of test score to the proportion of true non-masters falsely declared to be masters). A third index is the validity coefficient (the phi coefficient relating classifications based on test score and teacher judgment). Finally, indices of expected utility (the gain resulting from correctly identifying masters and non-masters), expected disutility (the loss incurred through errors in classification), and expected maximal utility (a weighted total of type one error probabilities) will be compared (Berk 1976).

All indices will be calculated for a range of possible cut-scores around the chosen standard. These computations will require MSDE to specify the relative severity of the two types of classification error and the relative utilities and disutilities of possible decisions. Constant utility and disutility functions will be used.

The predetermined cut-score, or any single cut-score, may not be optimal over all of the indices to be computed. The existing cut-score will be compared to other possible cut-scores. If another cut-score or scores is superior to the predetermined cut-score on one or more of the validity indices, this information can provide feedback and possible support for changing standards.

All of the cut-score validity procedures can occur before the actual bylaw is passed, therefore allowing even more information to be funneled into the judgmental process.

Conclusion

All standard-setting procedures will occur in the Spring and Summer of 1982. Each source of information will not have equal weight in the decision-making process. The item-based procedure will provide the most pertinent information.

When the steps outlined in this paper are completed, MSDE will be able to refer to a logical information gathering plan to support any final decisions about a passing score for the MFMT. Procedures such as these will also be used to set the standards for the three Maryland Life Skills Tests in Citizenship, Survival and World of Work.

ls

REFERENCES

Berk, R. A. Determination of optimal cutting scores in criterion-referenced measurement. Journal of Experimental Education, 45, 1976, 4-9.

Jaeger, Richard M. An Interactive Structured Judgment Process for Establishing Standards on Competency Tests: Theory and Application. Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, April 1981.

Popham, W. James. Modern Educational Measurement, Chapter 16, "Setting Standards." Prentice-Hall, Inc., Engelwood Cliffs, New Jersey: 1981, 371-399.

APPENDIX

Draft copy for Review

Dear Educator:

Beginning in 1983, Maryland students must pass the Maryland Functional Mathematics Test (MFMT) to be awarded a high school diploma. This year the Maryland State Board of Education will be establishing a passing standard for the test. This standard will be the minimum percent of items that a student must answer correctly to pass. In order to guide the Board in its deliberations, Maryland's educators, parents, and students are being invited to offer their opinions on what the passing standard on the test should be.

The accompanying materials describe the MFMT. Please review these materials. (More elaborate descriptions of the test's coverage can be found in Test Specifications for the Maryland Functional Mathematics Test, a copy of which has been furnished to your school's principal.) Then, considering the purpose of the MFMT, decide upon the minimum percent correct score that you think a student should be required to achieve in order to pass. Indicate this percent correct on the tear-off section at the bottom of this letter. Return the tear-off sheet to the appropriate person (location) designated in your school. All responses will be treated with total confidentiality.

Your preferences, along with those of other educators in Maryland, will be given serious consideration by the State Board. Thank you for your assistance.

Sincerely,

Appropriate Official
State Department of
Education

- -

EDUCATOR RESPONSE

I believe the passing standard for the Maryland Functional Mathematics Test should be ___ % correct.

SUMMARY REPORT FORM:

Directions: This report form is divided into three sections, one each for educators, parents, and students. For each section for which you have preference data to report, please provide the average passing standard, that is, the percent correct score recommended by that group of individuals. For example, if 23 parents returned tear-off slips with passing standards, add all their standards together and divide that total by 23 to determine the average standard they are recommending. Also please indicate the number of individuals who contributed to the recommended standard. In this instance, for example, there would have been 23 parents involved.

Please return the Summary Report Form in the stamped, self-addressed envelope which has been provided. Thank you for your assistance.

EDUCATORS

Recommended Passing Standard (average percent correct score): _____%

Number of Educators responding _____

PARENTS

Recommended Passing Standard (average percent correct score): _____%

Number of Parents responding _____

STUDENTS

Recommended Passing Standard (average percent correct score): _____%

Number of Students responding _____

DESCRIPTION OF THE MARYLAND FUNCTIONAL MATHEMATICS TEST

The purpose of the Maryland Functional Mathematics Test (MFMT), is to assure that Maryland high school graduates have met an acceptable standard of functional mathematics achievement.

The MFMT will be administered to ninth grade students each fall. Once passed, the test need not be re-taken. Students who do not pass at the first administration will be given at least three additional opportunities at later grade levels to pass equivalent forms of the test. All high school graduates will have to pass the test.

The test itself will consist of approximately 80 items similar to the samples given here. Students will be directed to round off their answers as needed. There is no time limit on the test.

The MFMT is organized around seven content areas. To provide a general idea of what the test covers, each of these seven areas is briefly described below, followed by a sample test item similar to those which will be included on the actual test. Correct answers are indicated by an asterisk(*). While the seven sample items do not reflect all types of items on the test, they do represent the average difficulty of the actual items on the test. This content description is based on a more complete set of content specifications (Test Specifications for the Maryland Functional Mathematics Test, Maryland State Department of Education, Baltimore, November, 1981).

Number Concepts

Students will be asked to match numbers to words or words to numbers, to rename fractions as percents, to rename percents as decimals, and to put groups of numbers in order from least to greatest.

Sample Test Item

Rename $\frac{3}{4}$ as a percent:

*(A) 75%

(B) $133\frac{1}{3}\%$

(C) $77\frac{3}{4}$

(D) .75%

*Indicates the correct answer

Whole Number Operations

Students will be asked to add, subtract, multiply, and divide whole numbers.

Sample Test Item

SUBTRACT:

$$\begin{array}{r} 6{,}843 \\ -1{,}966 \\ \hline \end{array}$$

(A) 5,123

(B) 4,887

*(C) 4,877

(D) 5,987

Mixed Number/Fraction Operations

Students will be asked to add and subtract mixed numbers (for example $2\ \frac{}{}$), to multiply whole numbers by fractions, and to determine missing terms in proportions.

Sample Test Item

ADD AND SIMPLIFY:

$$\begin{array}{r} 2\ \frac{1}{2} \\ +\ 3\ \frac{1}{3} \\ \hline \end{array}$$

(A) $5\ \frac{2}{5}$

(B) $5\ \frac{2}{3}$

*(C) $5\ \frac{5}{6}$

(D) $5\ \frac{1}{3}$

Decimal Operations

Students will be asked to add, subtract, multiply, and divide decimal fractions and to determine percentages.

Sample Test Item

MULTIPLY:

$$\begin{array}{r} 34.9 \\ \times\ .56 \\ \hline \end{array}$$

(A) 3.839

*(B) 19.544

(C) 195.44

(D) 38.39

*Indicates correct answer

Measurement

Students will be presented with pictured objects or word problems and asked to select appropriate measurement units, read pictured measuring instruments, calculate perimeters and areas, and determine elapsed time.

Sample Test Item

The height of a classroom door would best be measured in:

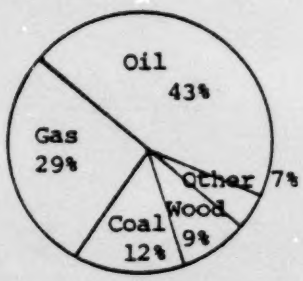
(A) cm
(B) km
*(C) m
(D) mm

Using Data

Students will be asked to answer questions requiring the proper interpretation of tables and graphs and to compute averages for sets of numbers.

Sample Test Item

ENERGY SOURCES FOR HOME HEATING



What total percent of the people use either wood or coal to heat their homes?

*(A) 21%
(B) 12%
(C) 16%
(D) 3%

*Indicates correct answer

Problem Solving

Students will be asked to solve word problems requiring one or two operations of addition, subtraction, multiplication, or division.

Sample Test Item

Car expenses for one month are:

Gasoline	$75.42
Battery	$35.00
Insurance	$43.00

What are the total expenses for the month?

(A) $143.42
*(B) $153.42
(C) $110.42
(D) $118.42

*Indicates correct answer

(This insert will go in half of all mail-out materials.)*

In the fall of 1981, the items which will be used in the MFMT were administered to representative samples of approximately 9,000 ninth-grade and 3,000 eleventh-grade Maryland students. The ninth-graders achieved an average of 61% correct; the eleventh-graders achieved an average of 72% correct. The table below gives you some more information about how students did on this test. It shows what percent of the ninth- and eleventh-grade students who took the test would have passed at different passing scores. So for example, if the passing score were 30% correct, 98% of the ninth-graders and 100% of the eleventh-graders would have passed. If the passing score were 90% correct, however, only 42% of the ninth-graders and 51% of the eleventh-graders would have passed.

	Students Who Would Pass	
Passing Score	Ninth-grade	Eleventh-grade
10%	100%	100%
30%	98%	100%
50%	84%	91%
70%	52%	62%
90%	42%	51%

*All figures in this section are hypothetical at this point. Data analysis will be completed and actual figures inserted before mailout.